Nonlinearities in the World Vegetable Oil Price System: El Niño Effects

David Ubilava
Department of Agricultural Economics
Purdue University

Matthew T. Holt
Department of Agricultural Economics
Purdue University

This Draft: April 30, 2009

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association 2009 AAEA & ACCI Joint Annual Meeting, Milwaukee, Wisconsin, July 26-29, 2009

Nonlinearities in the World Vegetable Oil Price System: El Niño Effects

David Ubilava[†] **Matthew T. Holt**[‡]

Abstract

In this research we estimate the effect of El Niño Southern Oscillation (ENSO) over time on market dynamics for eight major vegetable oil prices. We estimate a system for vegetable oil prices by using a smooth transition vector error correction model (STVECM) to analyze impacts of ENSO events on production, and, more interestingly, their asymmetric nature. The results of estimated Exponential STVECM and Quadratic STAR models, respectively for the system of oil price equations and the ENSO variable regressions, suggest a smooth transition between ENSO regimes, and provide a better overall fit to the data than do linear models. Effects of the ENSO shock are analyzed using generalized impulse-response functions (GIRFs). The nonlinear nature of these shocks is apparent, as the GIRFs are observed to be asymmetric depending on whether ENSO shocks are positive or negative. For most vegetable oil prices an ENSO shock has a permanent effect, meaning that prices advance to a new equilibrium level. Generally, a positive ENSO shock results in increased prices, and the opposite is true for the negative ENSO shock. The magnitude of the price change is largest for the coconut oil and palm kernel oil, and is the smallest for the ground nut oil. Also, it takes approximately two years for prices to stabilize at a new equilibrium level after the shock.

Keywords: Smooth Transition, Vector Error Correction, El Niño Southern Oscillation, Vegetable Oil Prices

[†]Department of Agricultural Economics, Purdue University, 403 W. State Street, West Lafayette, IN 47907–2056, USA. Telephone: 765-494-7709. Fax: 765-409-3336. E-mail: dubilava@purdue.edu.

[‡]Department of Agricultural Economics, Purdue University, 403 W. State Street, West Lafayette, IN 47907–2056, USA. Telephone: 765-494-7709. Fax: 765-494-9176. E-mail: mholt@purdue.edu.

1 Introduction

The role of weather conditions (climate anomalies) in agricultural commodity production and prices has long been of interest. Indeed, it is highly likely that for as long as mankind has engaged in the cultivation of crops for food and fiber consumption that there has been a keen recognition of the role of weather in food prices, production, and availability. For example, Lamb (1995) provides evidence of extreme price spikes for cereal grains in Northern and Western Europe during the worst years of the Little Ice Age, a period that was often associated with extremely poor growing conditions and encompassed effectively all of the 16th and 17th centuries. Likewise, Temin (2002) has linked prices for various commodities in ancient Babylon including, for example, barley and mustard, to seasonal growing conditions. And of course more recently various authors have examined the role of weather events and climate change on crop yields, land prices, and profitability. See, for example, Schlenker and Roberts (2006) for some comprehensive empirical work on the effects of weather on corn yields as well as a a recent literature review.

Aside from the obvious connections between rainfall and temperature on crop yields and prices, there is budding interest in the role of role of major climate anomalies on the performance of various economic variables. In part this interest stems from the growing recognition that even local growing conditions may be affected by large–scale climate events. To this end several studies have found statistically significant correlations between climatic events and economic behavior. See, for example, Handler and Handler (1983), Debelle and Stevens (1995), Solow et al. (1998), and Hansen et al. (1998). In recent years there has been specific attention paid to the relationship between the El Niño Southern Oscillation (ENSO) phenomena and commodity price movements as in, for example, Keppenne (1995), Letson and McCullough (2001), and Brunner (2002). In this research we estimate the effect of ENSO over time on market dynamics for eight major vegetable oil prices: soybean, sunflower seed, groundnut, coconut, palm kernel, palm, cottonseed, and rapeseed. While these oils are by no means perfect substitutes in consumption, they share many similar qualities, such as their use in food preperation, soap production, manufacturing of paints and medicines, and, more recently, in bio-fuel production. As a result there is a reasonably high degree of substitutability among these oils, the result being that vegetable oil prices are likely to be highly interlinked, at least, in the long run if not in the short run In and Inder (1997).

El Niño events are linked to oscillations of the ocean–atmosphere system in the tropical Pacific ocean, which in turn are identified with significant consequences for global weather conditions. In normal conditions the trade winds blow west across the tropical Pacific.

During El Niño events the trade winds weaken in the central and western Pacific leading to a depression of the thermocline in the eastern Pacific, and an elevation of the thermocline in the west, which results in anomalously warm sea surface temperatures interacting with the air above it in the eastern and central Pacific Ocean. Apparent consequences of ENSO events are increased rainfall across the southern tier of the U.S. and in Peru, and drought in the Western Pacific region. The counterpart of El Niño is La Niña, which, on the other hand, is associated with very intense trade winds, and colder–than–normal sea surface temperatures in the region. The whole El Niño – La Niña – El Niño cycle generally takes approximately four years, and may range between two and seven years. Taken together, this systematic fluctuation between El Niño – La Niña events defines the El Niño Southern Oscillation, or ENSO. These effects were first identified in the early 1920s by Sir Gilbert Thomas Walker (Walker, 1923), who eventually coined the phrase "Southern Oscillation."

There is substantial evidence that ENSO events have significant impacts on the production and prices for vegetable oils, and most notably for those produced in the Philippines and elsewhere in South Asia (e.g., palm oil and palm kernel oil). For example, a strong El Niño event typically results in drought in the Western Pacific region, which in turn curtails production of palm and palm kernel oil. In turn these price shocks will eventually impact the demands for close substitutes including soy oil, sunflower oil, and so forth. Therefore, there is reason to believe that ENSO events will have a an influence and perhaps a marked influence on market prices for vegetable oils, at least in the short run and perhaps even in the intermediate term. This later observation suggests that linear models, for example, standard Vector Autoregressions (VAR) or Vector Error Correction Models (VECM) may not be the first tools of choice to examine relationships between ENSO and vegetable oils prices; these models imply symmetric adjustments, that is, a positive shock (El Niño) is the mirror image of a negative shock (La Niña). The implication is that models which allow for the possibility of asymmetric adjustments may be preferred.

It is also reasonable to conjecture that responses to ENSO shocks are asymmetric–a strong El Niño event may result in price dynamics for vegetable oils that are very different than, say, those associated with a strong La Niña event. This later observation suggests that linear models, for example, standard Vector Autoregressions (VAR) or Vector Error Correction Models (VECM) may not be the first tools of choice to examine relationships between ENSO and vegetable oils prices; these models imply symmetric adjustments, that is, a positive shock (El Niño) is the mirror image of a negative shock (La Niña). The implication is that models which allow for the possibility of asymmetric adjustments may be preferred. We propose to examine these relationships in considerable detail here.

We estimate the system of vegetable oil price time series using the smooth transition autoregressive version of the vector error correction model. Results of specification tests support an exponential STVECM and a quadratic STAR model for the system of oil price equations and the ENSO regressions, respectively. Moreover, the model estimates for the STVECM do indeed suggest a smooth transition between regimes. As well, the non–linear models provide a better overall fit to the data than do linear models. Effects of the ENSO shock are analyzed using generalized impulse-response functions (GIRFs). The non–linear nature of these shocks is apparent, as the GIRFs are observed to be asymmetric depending on whether ENSO shocks are positive or negative. For most vegetable oil prices an ENSO shock has a permanent effect, meaning that prices advance to a new equilibrium level. Generally, a positive ENSO shock results in increased prices while the opposite is true for a negative shock. The magnitude of the price change is largest for coconut and palm kernel oil, and is smallest for ground nut oil. Also, it takes approximately two years for prices to stabilize at a new equilibrium level after the shock.

2 Econometric Framework

In this section we outline the econometric approach used here to investigate the potential relationship between ENSO events and commodity prices, as well as the possible asymmetry that exists between the sign/magnitude of these shocks and their resulting price impacts. We begin by focusing on linear modeling techniques for systems of equations, most notably, VARs and VECMs. We then turn our attention to the family of smooth transition models including smooth transition vector error correction models (STVECMs).

2.1 Vector Error Correction Model

The n-equation *vector autoregressive* (VAR) model of order p and can be expressed as:

$$X_t = \sum_{i=1}^{p} \Pi_i X_{t-i} + \mu + \Phi Z_t + \epsilon_t, \qquad t = 1, \ldots, T \tag{1}$$

where X_{t-i} $(i = 0, \ldots, p)$ are $n \times 1$ vectors of dependent variables; Z_t is a $m \times 1$ vector of exogenous variables and/or seasonal dummies, and ϵ_t is a $n \times 1$ vector of white noise process with positive definite covariance matrix Σ. Π_i are $n \times n$ matrices, μ is a $n \times 1$ vector, and Φ is a $n \times m$ matrix of parameters to be estimated. This also can be rewritten in a first-difference form as:

$$\Delta X_t = \sum_{i}^{p-1} \Gamma_i \Delta X_{t-i} + \Pi X_{t-1} + \mu + \Phi Z_t + \epsilon_t \tag{2}$$

where

$$\begin{aligned} \Gamma_i &= -\left(\Pi_{i+1} + \ldots + \Pi_p\right), \qquad i = 1, \ldots, p-1 \\ \Pi &= -\left(I - \Pi_1 - \ldots - \Pi_p\right) \end{aligned}$$

Here Π is a matrix of coefficients that potentially contains the information about the relationships between the variables in the system. Depending on a rank of Π matrix there may be three possible outcomes: 1) $Rank\,(\Pi) = n$, indicating that Π is a full-rank matrix, and the vector process x_t is stationary; 2) $Rank\,(\Pi) = 0$, indicating that Π is a null matrix, and Equation 2 is a traditional differenced VAR model; and 3) $Rank\,(\Pi) = r$, $0 < r < n$, indicating that there are $n \times r$ matrices α and β, such that $\alpha\beta' = \Pi$, where β is a matrix of cointegrating vectors, such that $\beta' X_t$ is a stationary process, even though X_t itself is not; and α is a matrix of speed-of-adjustment parameters. Johansen's test (see Johansen, 1988; Johansen and Juselius, 1990) is used to find the rank of Π matrix.

Johansen's test is performed on a system of VAR(p) equations. The rank, r, of the Π matrix is revealed using λ_{max} and λ_{trace} statistics, which are calculated as follows:

$$\lambda_{max}(r, r+1) = -T \ln\left(1 - \hat{\lambda}_{r+1}\right) \tag{3}$$

$$\lambda_{trace}(r) = -T \sum_{i=r+1}^{m} \ln\left(1 - \hat{\lambda}_i\right) \tag{4}$$

where T is the number of usable observations; r is the rank of the Π matrix; n is the number of equations; and $\hat{\lambda}_i$ are the eigenvalues of the estimated Π. Further, the hypothesis of an intercept in the cointegrating vector(s) may be tested using the following statistic:

$$\xi_{n-r} = -T \sum_{i=r+1}^{m} \ln\left(\frac{1 - \hat{\lambda}_i^*}{1 - \hat{\lambda}_i}\right) \tag{5}$$

where $\hat{\lambda}_i^*$ and $\hat{\lambda}_i$ are the eigenvalues of the estimated restricted and unrestricted Π matrices, respectively, where the restriction implies the presence of the intercept in the cointegrating vectors. Test statistic, ξ, is chi-square distributed with $(n-r)$ degrees of freedom.

Assuming that the rank of the Π matrix is r, and an intercept is included in the cointegrating vector(s), the VAR model can be rewritten in a *vector error correction* (VEC) form as follows:

$$\Delta X_t = \sum_{i}^{p-1} \Gamma_i \Delta X_{t-i} + \upsilon' \hat{e}_{t-1} + \Phi Z_t + \epsilon_t \tag{6}$$

where $e_{t-1} = \tilde{\beta}' \tilde{X}_{t-1}$ is a vector of estimated error-correction terms, where $\tilde{\beta}$ is a $(n+1) \times r$ matrix of normalized cointegrating vectors including the intercept, obtained from the Johansen's procedure, and $\tilde{X}_{t-1} = \left(X'_{t-1}, 1\right)'$; and υ is a $r \times 1$ vector of parameters to be estimated.

2.2 Smooth Transition Models

Class of smooth transition autoregressive (STAR) models is widely used in the studies attempting to model the asymmetric cyclical variations and turbulent periods (e.g. Terasvirta and Anderson, 1992; Terasvirta, 1995; Hall et al., 2001). STAR model of order p can be specified as:

$$\Delta y_t = \phi_1' x_t \left(1 - G\left(s_t; \gamma, c\right)\right) + \phi_2' x_t G\left(s_t; \gamma, c\right) + \epsilon_t \tag{7}$$

or, alternatively:

$$\Delta y_t = \phi_1' x_t + \left(\phi_2 - \phi_1\right)' x_t G\left(s_t; \gamma, c\right) + \epsilon_t \tag{8}$$

where y_t is a dependent variable; $x_t = (1, \Delta y_{t-1}, \ldots, \Delta y_{t-p}, y_{t-1}, Z_t)'$, is a matrix of independent variables, where $Z_t = (Z_{1,t}, \ldots, Z_{m,t})$ is a vector of exogenous or seasonal dummy variables; Δ is the first difference operator; $\phi_1 = (\phi_{1,0}, \ldots, \phi_{1,k})$, and $\phi_2 = (\phi_{2,0}, \ldots, \phi_{2,k})$ are parameter vectors; and $\epsilon_t \sim N\left(0, \sigma_\epsilon^2\right)$.

This is a two-regime model (which may be extended to any k-regime model in a trivial manner), where $G\left(s_t; \gamma, c\right)$ is a transition function, by construction bounded between zero and one, and where s_t is a transition variable, and γ and c are, respectively, smoothness and location parameters to be estimated. Often the lagged difference $\Delta y_{t-d} = s_t = (y_{t-d} - y_{t-d-1})$ or lagged seasonal difference $\Delta_J y_{t-d} = s_t = (y_{t-d} - y_{t-d-J})$ variable is used as a transition variable, where d is referred as the delay parameter. Alternatively, $t^* = (t/T)$, where T is the length of the time series, may be used as a transition variable leading to the time-varying autoregressive (TVAR) model.

In empirical studies the logistic and exponential transition functions are used most frequently, respectively forming logistic STAR (LSTAR) and exponential STAR (ESTAR) models. Another frequently used functional form is quadratic STAR (QSTAR). These transition functions are defined as follows:

$$G^L(s_t;\gamma,c) = \left(1+\exp\left[-\gamma\left(s_t-c\right)\right]\right)^{-1} \tag{9}$$

$$G^E(s_t;\gamma,c) = 1-\exp\left[-\gamma\left(s_t-c\right)^2\right] \tag{10}$$

$$G^Q(s_t;\gamma,c) = \left(1+\exp\left[-\gamma\left(s_t-c_1\right)\left(s_t-c_2\right)\right]\right)^{-1} \tag{11}$$

In a transition function γ is a non-negative parameter. LSTAR model approaches a linear AR model when $\gamma \to 0$, and a threshold autoregressive model (TAR) when $\gamma \to \infty$. On the other hand, ESTAR approaches linear AR in both cases, when $\gamma \to 0$ and $\gamma \to \infty$. QSTAR approaches to linear AR model when $\gamma \to 0$, and a TAR model with two threshold variables when $\gamma \to \infty$.

A family of smooth transition autoregressive models can be extended to a multivariate case, forming a multivariate STAR (MSTAR) model. One specific case of the MSTAR model is a smooth transition vector error correction (STVEC) model, expressed as follows:

$$\Delta X_t = \Psi_1 W\left(1-G\left(s_t;\gamma,c\right)\right) + \Psi_2 W G\left(s_t;\gamma,c\right) + \epsilon_t \tag{12}$$

or, alternatively

$$\Delta X_t = \Psi_1 W + \left(\Psi_2-\Psi_1\right) W G\left(s_t;\gamma,c\right) + \epsilon_t \tag{13}$$

where $W = (X_{t-1}, \ldots, X_{t-p+1}, \hat{e}_{t-1}, Z_t')'$; and $\Psi_k = (\Gamma_{k,1}, \ldots, \Gamma_{k,(p-1)}, \upsilon_k', \Phi_k)$ are parameters to be estimated. $G\left(s_t;\gamma,c\right)$, for any given s_t, γ and c may be restricted to be the same, or vary, across the equations, respectively forming one common, or n unique transition functions for the equations.

The nonlinearity and parameter constancy in the smooth transition autoregressive model is tested using the auxiliary regression, since the conventional test statistics are not directly applicable to the STAR-type models due to the identification problems (see Luukkonen et al., 1988)[1]. Suitable auxiliary regressions are first- or third-order Taylor series expansion of the

[1]Here we only present the brief overview of the test procedure. For more details refer to Luukkonen et al.

transition functions. The third-order Taylor series expansion is a preferred form of the test, because it embeds tests against LSTAR and ESTAR/QSTAR, along with the general test against nonlinearity. Testable auxiliary equation, in general terms, may be expressed as follows:

$$\Delta y_t = \theta_0' x_t + \theta_1' x_t s_t + \theta_2' x_t s_t^2 + \theta_3' x_t s_t^3 + \nu_t \tag{14}$$

where θ_s, $s = 0, \ldots 3$ are vectors of parameters to be estimated, and $\nu_t \sim IID\,(0, \sigma_\nu^2)$. The underlying hypotheses are: $H_0 : \theta_1 = \theta_2 = \theta_3 = 0$ for no nonlinearity, $H_{03} : \theta_3 = 0$ and $H_{01} : \theta_1 = 0 | \theta_2 = \theta_3 = 0$ for no logistic nonlinearity, and $H_{02} : \theta_1 = \theta_2 = 0 | \theta_3 = 0$ for no exponential (or quadratic) nonlinearity. The F version of the LM test statistic is preferred for the small and moderate samples:

$$LM = \frac{(SSR_0 - SSR_1)\,/\,(lq)}{SSR_1/\,(T - (m+1)q)} \tag{15}$$

where SSR_0 and SSR_1 are sums of squared residuals of the restricted and unrestricted models, respectively; q is the number columns in the matrix of independent variables; l takes values 1 and 3, and m takes values 1, 2 and 3, depending on a hypothesis being tested; T is the total size of the time series. This approach can be used for a single equation autoregressive model, or in case when each equation in the system has its unique transition function. However, for a system of equations with identical transition function across the equations, the system of auxiliary equations will be as follows:

$$\Delta X_t = \Theta_0' W_t + \Theta_1' W_t S_t + \Theta_2' W_t S_t^2 + \Theta_3' W_t S_t^3 + \nu_t \tag{16}$$

where Θ_S, $S = 0, \ldots 3$ are matrices of parameters to be estimated; and S_t is a $n \times 1$ vector of the transition variable; other variables are as defined above. For testing the hypotheses similar to those defined above, we should apply the likelihood ratio test statistics of the form:

$$LR = T(\ln|\Sigma_u| - \ln|\Sigma_r|) \tag{17}$$

which is χ^2 distributed with degrees o freedom equal to the number of restrictions; and where Σ_u and Σ_r are estimated variance-covariance matrices of the error terms from the

(1988); Terasvirta and Anderson (1992); Eitrheim and Teräsvirta (1996) unrestricted and restricted models, respectively (see also Weise, 1999).

2.3 Generalized Impulse Response Functions

Generalized impulse-response (GI) function, proposed by Koop et al. (1996), is defined as follows:

$$GI\left(h, \nu, \omega_{t-1}\right) = E\left(y_{t+h} | \nu, \omega_{t-1}\right) - E\left(y_{t+h} | \omega_{t-1}\right) \tag{18}$$

where h is a horizon of the shock effect, ν is a shock in period t, and ω_{t-1} is a history. Assuming that ν and ω_{t-1} are random variables, so that $\nu \in N$ and $\omega_{t-1} \in \Omega_{t-1}$, GI function itself represents the random variable realization:

$$GI\left(h, N, \Omega_{t-1}\right) = E\left(y_{t+h} | N, \Omega_{t-1}\right) - E\left(y_{t+h} | \Omega_{t-1}\right) \tag{19}$$

Further, it is possible to define GI function as a random variable in terms of the history, conditional on a particular shock:

$$GI\left(h, \nu, \Omega_{t-1}\right) = E\left(y_{t+h} | \nu, \Omega_{t-1}\right) - E\left(y_{t+h} | \Omega_{t-1}\right) \tag{20}$$

or as a random variable in terms of the shock, conditional on a particular history:

$$GI\left(h, N, \omega_{t-1}\right) = E\left(y_{t+h} | N, \omega_{t-1}\right) - E\left(y_{t+h} | \omega_{t-1}\right) \tag{21}$$

and, eventually, as a random variable in terms of subsets of shocks and histories:

$$GI\left(h, \Re, \Im_{t-1}\right) = E\left(y_{t+h} | \Re \in N, \Im_{t-1} \in \Omega_{t-1}\right) - E\left(y_{t+h} | \Im_{t-1} \in \Omega_{t-1}\right) \tag{22}$$

Below we present the approach of estimating GI based on the latter condition. A stocastic simulation approach is used to obtain the estimates of the response function. Given a particular subsets of histories and shocks, the GI function can be defined as:

$$GI\left(h, \Re, \Im_{t-1}\right) = 1/J \sum_{j=1}^{J} GI^{j}\left(h, \Re, \omega_{t-1}^{j}\right) \tag{23}$$

where

$$GI^j\left(h, \Re, \omega_{t-1}^j\right) = 1/D \sum_{d=1}^{D} GI^{d,j}\left(h, \nu^d, \omega_{t-1}^j\right)$$

and where

$$GI^{d,j}\left(h, \nu^d, \omega_{t-1}^j\right) = E\left(y_{t+h} | \nu^d, \omega_{t-1}^j\right) - E\left(y_{t+h} | \omega_{t-1}^j\right)$$

The expected realizations of y_{t+h} are obtained using bootstrapping procedure, so that:

$$E\left(y_{t+h}\right) = 1/R \sum_{r=1}^{R} y_{t+h}^r$$

where

$$y_{t+h}^r = y\left(h, \nu^d, \omega_{t+h-1}^j, \theta, \epsilon^r\right)$$

where θ are estimated parameters from the model, ϵ^r are R randomly drawn shocks, D and J are total number of selected shocks and drawn histories, respectively.

3 Empirical Framework

3.1 Data

In this research we use monthly time series of ENSO anomaly and vegetable oil prices covering the period between January 1972 and December 2005. Anomaly of the *Niño 3.4* monthly average index as an ENSO measure is derived from the index tabulated by the Climate Prediction Center at the U.S. National Oceanic and Atmospheric Administration. This index measures the difference in Sea Surface Temperature (SST) in the area of the Pacific Ocean between 5^oN-5^oS and 170^oW-120^oW, and thus is a strong indicator of ENSO occurrence. *Niño 3.4* monthly measure is an average of daily values interpolated from the weekly measures obtained both from satellites and actual locations around the Pacific. The anomaly is the deviation of the *Niño 3.4* monthly measure from the average historic measure for that particular month from the period 1971-2000. The other widely used ENSO measure (also used by most of the related aforementioned studies) is Southern Oscillation Index (SOI), a measure derived from the sea level pressure. In current research SST was chosen over SOI essentially because the former has a stronger relationship with vegetable oil prices (Holt and Inoue, 2006).

Vegetable oil price data was obtained from ISTA Mielke GmbH, better known as OIL WORLD. All prices are in U.S. dollars per tonne, and are either FOB or CIF[2]. Plotting the vegetable oil prices together allows us to see their apparent co-movement over the time (see Figure 1)[3]. It is clearly seen that prices in general move together, and occasional divergences are soon followed by convergences. It is, further, straightforward to imagine that all [other pairs of] the oil prices co-move at least in a long run.

For the estimation purposes we deflated the nominal prices using the PPI for commodities obtained from the U.S. Bureau of Labor Statistics. Further, real prices were transformed to the natural log form, so that changes are explained in percentage terms. From here forward whenever vegetable oil price is mentioned it is to be considered as real price in natural logarithmic form, unless otherwise stated.

3.2 Estimation

Let $X_t = (soy_t, sun_t, grn_t, coc_t, plk_t, plm_t, ctn_t, rap_t)'$ be a vector of time series, where dependent variables are natural logs of the real vegetable prices stacked in a sequence as listed in the introduction. Then the associated VAR model is specified as in Equation 1, where exogenous variable $Z_t = (ENSO_t, ENSO_{t-1}, \ldots, ENSO_{t-m})'$ are ENSO in levels[4]. Using AIC number of lags in the VAR, p, was set to 4, and m was set to 2. Further we applied Johansen's procedure to test the hypothesis of cointegration. λ_{max} statistics supported two cointegrating vectors, and λ_{trace} statistics supported four cointegrating vectors hypotheses at $\alpha = 0.05$ significance level (see Table 1). This type of outcome is not unusual in empirics, and in such occasions λ_{max} statistics is preferred to λ_{trace}, because the former has the sharper alternative hypothesis (e.g. Enders, 2004, p. 354). The hypothesis of an intercept in the cointegrating vectors (Equation 5) was not rejected. Therefore the obtained VECM is specified as in Equation 6.

With the obtained specification VECM requires estimating $((n \times n) \times (p-1) + r \times n +$

[2]FOB - Free on Board; CIF - Cost, Insurance, and Freight

[3]It would be visually inconvenient to present the time series of all eight vegetable oil prices on the same graph. Instead, we only plotted two the most diverging oil prices, selected by maximizing the sum of squared differences across the pairs of normalized vegetable oil prices.

[4]The joint hypothesis of presence of monthly seasonal dummies in the model was rejected. In fact, only 6 out of total 88 parameters of the monthly dummy variables were statistically significant at $\alpha = 0.1$ level. With an anticipation of losing degrees of freedom due to estimating the smooth transition type of model, we did not include monthly dummies in the model to retain some estimation flexibility

$m \times n) = 232$ parameters, which in a smooth transition VECM application will reach 466 (twice the amount of VECM parameters plus two additional for the associated transition function). Often it is the case that number of these parameters are far from being statistically significant, and their presence in the model may even distort the precision of statistical inference (Bruggemann and Lutkepohl, 2001). We, therefore, apply a strategy proposed by Bruggemann and Lutkepohl (2001) to specify the *subset VECM*. Namely, each of the n VECM equations is estimated using OLS, and regressors with the smallest absolute t-value are deleted sequentially until the smallest remaining t-value is not smaller than some threshold τ, where $\tau_i = \sqrt{(\exp(c_\tau/T) - 1)(T - k + i - 1)}$, and where T is the effective sample size, k is the number of parameters in the unrestricted model, and i is the iteration number; c_τ is to be set so that provided statistics will represent an information criterion of choice. In this research c_τ is set to 2, and hence, the AIC is used for selection the parameters in each equation. It is notable, that parameters of the error-correction terms were restricted to be present in each equation. The underlying reasoning is that it may be the case that in a smooth transition form of the VECM, a parameter associated with a certain error correction term may turn out to be significant in one (or both) regimes, even if it was not significant in a linear VECM. After eliminating the parameters following the above-described strategy, only 132 parameters were retained to be estimated in the VECM.

The ENSO is assumed to be weakly exogenous variable in the model. That is, vegetable oil prices are contemporaneously correlated with ENSO, but not the other way around. With this assumption holding, ENSO equation can be independently estimated in an univariate mode. The optimal number of lags in ENSO equation was set to be equal to 4, using AIC, and augmented Dickey-Fuller (ADF) statistics rejected the unit root hypothesis at $\alpha = 0.01$ level.

Taking all the aforementioned into the account, the subset VECM with the exogenous variable will have the following form:

$$\begin{aligned} \Delta x_t &= \sum_{i=1}^{p-1} \Gamma_i \Delta x_{t-i} + \upsilon' \hat{e}_{t-1} + \sum_{j=0}^{m} \Phi_j ENSO_{t-j} + \epsilon_t \\ \Delta ENSO_t &= \mu + \sum_{i=1}^{q-1} \beta_i \Delta ENSO_{t-i} + \beta_l ENSO_{t-1} + \eta_t \end{aligned} \tag{24}$$

where

$$\begin{bmatrix} \epsilon_t \\ \eta_t \end{bmatrix} \sim \mathrm{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma_\epsilon & 0 \\ 0 & \sigma_\eta^2 \end{bmatrix} \right)$$

and where variables and parameters are as defined above, and p, m and q are 4, 2, and 4, respectively. Note that because we work with the subset VECM, some of the elements of Γ_i and Φ_j are restricted to zero.

The following step of the estimation is to specify the suitable STAR and STVEC models for ENSO and the system of vegetable oil prices, respectively. For this, we apply the auxiliary regressions as presented in Equation 16, and the adequate hypotheses are tested using Equations 15 and 17, respectively, for the univariate and multivariate cases. We used up to 6 lags of ENSO and seasonally differenced ENSO as the candidate transition variables. Based on these test results (see Table 2), and also after comparing the log-likelihood functions and AIC of different potential transition functions[5] the following form of STVECM was selected:

$$\begin{aligned} \Delta X_t &= \Psi_1 W_t + (\Psi_2 - \Psi_1) W_t G(s_t; \gamma, c) + \epsilon_t \\ \Delta y_t &= \phi_1' w_t + (\phi_2 - \phi_1)' w_t F(s_t; \gamma, c) + \eta_t \end{aligned} \tag{25}$$

where Y_t is a vector of vegetable oil prices, and y_t is ENSO; $W_t = (\Delta X_{t-1}, \ldots, \Delta X_{t-3}, \hat{e}_{1,t-1}, \hat{e}_{2,t-1}, y_t, \ldots, y_{t-2})'$ and $w_t = (1, \Delta y_{t-1}, \ldots, \Delta y_{t-3}, y_{t-1})'$; and $\Psi_k = (\Gamma_{k,1}, \ldots, \Gamma_{k,3}, \upsilon_k', \Phi_k)$ and $\phi_k = (\alpha_k, \beta_{k,1}, \ldots, \beta_{k,3}, \beta_{k,l})'$, $k = 1, 2$; and where transition functions, $G(s_t; \gamma, c)$ and $F(s_t; \gamma, c)$ are exponential and quadratic transition functions, respectively, where $s_t^G = ENSO_{t-6}$ and $s_t^F = \Delta_{12} ENSO_{t-1}$.

The resulted transition functions are plotted in Figures 2 and 3.

3.3 Simulations

To analyze the effect of the ENSO shock on vegetable oil prices we use a generalized impulse-response (GI) approach. We plotted two transition variables used in the STVECM estimation to locate the candidate subsets of histories (see Figure 4).

[5] For example, in case of ENSO equation, the test for nonlinearity suggested LSTAR model with the first lag of the seasonally differenced variable as the transition variable, but eventually the QSTAR model, with the same transition variable, was selected for the estimation, because it yielded larger log-likelihood value and smaller AIC compared to LSTAR, or ESTAR with the fifth lag of the seasonally differenced variable as the transition variable. Analogously, in case of the system of vegetable oil prices, the first lag of the seasonally differenced ENSO was the first choice for the transition variable, but the hypothesis against the ESTAR model with first lag of ENSO was rejected with the stronger statistical significance.

We selected three most interesting subsets for the GI analysis: 1) normal regime while the seasonal change of the SST is close to zero (NN on the graph); 2) El Niño regime while the seasonal change of the SST is positive (EE); and 3) La Niña regime while the seasonal change of the SST is negative (LL). That is, we intend to simulate the behavior of vegetable oil prices with respect to the positive and negative shocks of ENSO (SST), given that initial conditions are in normal or extreme modes.

Using the procedure formulated above, we compute GI functions in the following manner. All 9 histories of the LL regime, and random samples of 9 histories from the NN and EE regimes are drawn to initialize the starting values from the data used to estimate the model. Values of the normalized initial shock are set equal to $\nu = \pm 0.1\hat{\sigma}_\eta, \pm 0.2\hat{\sigma}_\eta, \ldots, \pm 1.9\hat{\sigma}_\eta$, where $\hat{\sigma}_\eta$ is the estimated standard deviation of the residuals from the ENSO STAR model. The maximum forecast horizon is set at 48 (4 years). An analytical expression for the conditional expectation of GI function (Equations 18-22) are not available for $h > 0$, therefore, expectations are evaluated numerically using 512 bootstrap simulations. Under the assumption of weak exogeneity of ENSO variable, for each iteration, in the first stage ENSO forecasts are obtained with and without the initial shock, and by using STAR residuals sampled with replacement; then, in the second stage, the obtained forecasts of the ENSO variable and STVEC normalized residuals sampled with replacement are used to obtain the forecasts for the system of vegetable oil prices. With 19 shocks, 9 histories and 512 iterations, total of 87,552 GI vectors of length 48 for each, positive and negative shock, are calculated. The expected GI functions are obtained using the Equation 2.3. Finally, by totaling the obtained GI functions for the first differences, the impulse responses for the levels of vegetable oil prices are constructed as follows:

$$GI_X\left(h, \nu, \omega_{t-1}\right) = \sum_{i=0}^{h} GI_{\Delta X}\left(h, \nu, \omega_{t-1}\right) \tag{26}$$

Expected GI functions of the NN, EE and LL regimes are plotted in Figure 5.

4 Results and Discussion

The values of the log-likelihood functions and AIC of the ENSO and system of oil prices regressions for linear and non-linear models are presented in Table 3. The non-linear models fit better both, the system of equations and the exogenous variable regressions.

A “normal” regime for the ENSO STAR model is estimated to be between -3.37 and 1.24

of the seasonal SST change. That also means that if the seasonal change is less than -3.37 then the model will predict the switch to the El Niño (extreme) regime; alternatively, if the seasonal change is more than 1.24 then the model will predict the switch to the La Niña (extreme) regime. For the system of vegetable oil prices, "normal" regime is estimated to range between -0.55 and 0.23 of the SST. This result fits well with the formal definition of the ENSO anomaly, where the normal condition is when SST index is in the neighborhood of zero. In both non-linear regressions the transition from one regime to another is happening smoothly, with the transition parameters being 1.33 and 4.17 for the STAR and STVEC models, respectively.

Effects of the ENSO shock are analyzed using GI functions. The non-linear nature of these shocks are easily noticed, as the GI functions tend to be asymmetric for positive and negative shocks (see Figure 5). For most of the vegetable oil prices the ENSO shock has a permanent effect, meaning that prices do not return to their initial, before the shock levels. Generally, positive ENSO shock results in increased vegetable oil prices, and the opposite is true for the negative ENSO shock. That is, in general higher prices are associated with the El Niño regime, and lower prices are associated with the La Niña regime. The magnitude of the price change is larger for the coconut oil and palm kernel oil prices, and is the lowest for the ground nut oil prices. Also, it takes approximately two-year period for prices to stabilize at a certain level, after the shock.

Further, in all cases, the positive shock has a larger magnitude when the initial conditions are LL compared to EE; alternatively, the negative shocks have a larger magnitude (in absolute terms) when the initial conditions are EE compared to LL. When initial conditions are NN, in all but two cases, the positive ENSO shock has a larger effect on prices, compared to the EE and LL conditions; and the negative ENSO shock has a smaller effect on prices compared to the EE condition, but is mostly similar to the LL initial condition. This means that, for example, if the ENSO anomaly is in the La Niña regime, then further deviation towards the La Niña will result in reduction of prices at larger extent, compared to the cases when the ENSO anomaly is in the normal or El Niño regimes. On the other hand, if the ENSO anomaly is in the El Niño regime, then further deviation towards the El Niño will result in increase of the vegetable oil prices at larger extent, compared to the case when the ENSO anomaly is in the La Niña regime, but at smaller extent, if the ENSO anomaly is in the normal regime.

5 Conclusions

In this paper we attempted to explore a nonlinear nature of the vegetable oil price dynamics with respect to the ENSO anomaly. Results of the research revealed the nonlinear behavior of the prices conditional on the state of nature and the direction of shocks of the ENSO anomaly. Overall, the negative deviation from the normal ENSO regime - La Niña results in lower vegetable oil prices, and the positive deviation - El Niño results in higher vegetable oil prices. Some oils are more responsive to the ENSO shock (e.g. coconut oil and palm kernel oil) than others (e.g. ground nut oil and cotton seed oil). The magnitude of the effect of the positive (negative) ENSO shock on vegetable oil prices is larger, if the ENSO anomaly is in the El Niño (La Niña) regime, and smaller if the ENSO anomaly is in the La Niña (El Niño) regime.

Arguably strong cointegration between the vegetable oil prices, and their impact on number of socio-economic factors in the world of oil exporters and importers, coupled with the fact that the behavior of these prices may be explained by an exogenous variable such as ENSO, makes this research urgent from the perspective of the policy implications.

References

Bruggemann, R. and H. Lutkepohl (2001). Lag selection in subset VAR models with an application to a US monetary system. *Econometric Studies*.

Brunner, A. (2002). El Nino and World Primary Commodity Prices: Warm Water or Hot Air? *Review of Economics and Statistics 84*(1), 176–183.

Debelle, G. and G. Stevens (1995). *Monetary Policy Goals for Inflation in Australia*. Economic Research Dept., Reserve Bank of Australia.

Eitrheim, Ø. and T. Teräsvirta (1996). Testing the Adequacy of Smooth Transition Autoregressive Models. *Journal of Econometrics 74*(1), 59–75.

Enders, W. (2004). *Applied Econometric Time Series*. Wiley.

Hall, A., J. Skalin, and T. Teräsvirta (2001). A Nonlinear Time Series Model of El Nińo. *Environmental Modelling and Software 16*(2), 139–146.

Handler, P. and E. Handler (1983). Climatic Anomalies in the Tropical Pacific Ocean and Corn Yields in the United States. *Science 220*(4602), 1155–1156.

Hansen, J., A. Hodges, and J. Jones (1998). ENSO Influences on Agriculture in the Southeastern United States*. *Journal of Climate 11*(3), 404–411.

Holt, M. and A. Inoue (2006). Climate Anomalies and World Primary Commodity Prices: The Effects of El Nino and His Primos Segundos as Viewed Through a Rolling Window. NCSU Working Paper. Draft: Apr. 10, 2006.

In, F. and B. Inder (1997). Long-run Relationships Between World Vegetable Oil Prices. *The Australian Journal of Agricultural and Resource Economics 41*(4), 455–470.

Johansen, S. (1988). Statistical Analysis of Cointegration Vectors. *Journal of Economic Dynamics and Control 12*(2/3), 231–254.

Johansen, S. and K. Juselius (1990). Maximum Likelihood Estimation and Inference on Cointegration with Applications to the Demand for Money. *Oxford Bulletin of Economics and Statistics 52*(2), 169–210.

Keppenne, C. (1995). An ENSO Signal in Soybean Futures Prices. *Journal of Climate 8*(6), 1685–1689.

Koop, G., M. Pesaran, and S. Potter (1996). Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics 74*(1), 119–147.

Lamb, H. (1995). *Climate, History, and the Modern World, 2nd ed.* Routledge.

Letson, D. and B. McCullough (2001). ENSO and Soybean Prices: Correlation without Causality. *Journal of Agricultural and Applied Economics 33*(3), 513–522.

Luukkonen, R., P. Saikkonen, and T. Terasvirta (1988). Testing Linearity Against Smooth Transition Autoregressive Models. *Biometrika 75*(3), 491–499.

Schlenker, W. and M. Roberts (2006). Nonlinear Effects of Weather on Corn Yields. *Review of Agricultural Economics 28*, 391–398.

Solow, A., R. Adams, K. Bryant, D. Legler, J. O'Brien, B. McCarl, W. Nayda, and R. Weiher (1998). The Value of Improved ENSO Prediction to US Agriculture. *Climatic Change 39*(1), 47–60.

Temin, P. (2002). Price Behavior in Ancient Babylon. *Explorations in Economic History 39*, 46–60.

Terasvirta, T. (1995). Modelling Nonlinearity in US Gross National Product 1889–1987. *Empirical Economics 20*(4), 577–597.

Terasvirta, T. and H. Anderson (1992). Characterizing Nonlinearities in Business Cycles using Smooth Transition Autoregressive Models. *Journal of Applied Econometrics 7*, S119–S136.

Walker, G. (1923). Correlation in seasonal variations of weather. viii. a preliminary study of world–weather. *Memoirs of the Indian Meteorological Department 24*(4), 75–131.

Weise, C. (1999). The Asymmetric Effects of Monetary Policy: A Nonlinear Vector Autoregression Approach. *Journal of Money Credit and Banking 31*, 85–108.

Tables

	λ_{max}	5%	λ_{trace}	5%
$r \leq 7$	3.10	9.24	3.10	9.24
$r \leq 6$	4.97	15.67	8.08	19.96
$r \leq 5$	19.40	22.00	27.47	34.91
$r \leq 4$	22.21	28.14	49.69	53.12
$r \leq 3$	31.63	34.40	81.32	76.07
$r \leq 2$	40.07	40.30	121.39	102.14
$r \leq 1$	54.78	46.45	176.17	131.70
$r = 0$	61.51	52.00	237.68	165.58

Table 1: λ_{max} and λ_{trace} Test Statistics and Critical Values

Nonlinearities in ENSO

	LM	LM_3	LM_2	LM_1	Model
$\Delta_{12}y_{t-1}$	0.022	0.019	0.255	0.155	LSTAR
$\Delta_{12}y_{t-4}$	0.068	0.098	0.031	0.836	ESTAR/QSTAR
$\Delta_{12}y_{t-5}$	0.043	0.327	0.016	0.321	ESTAR/QSTAR

Nonlinearities in Vegetable Oil Prices

	LR	LR_3	LR_2	LR_1	Model
y_{t-6}	0.013	0.061	0.000	0.989	ESTAR/QSTAR
$\Delta_{12}y_{t-1}$	0.000	0.039	0.014	0.027	ESTAR/QSTAR
$\Delta_{12}y_{t-6}$	0.003	0.027	0.001	0.690	ESTAR/QSTAR

Numbers are p-values of the associated test statistics

Table 2: Candidate Transition Variables for STAR and STVEC models

	$ENSO_{AR}$	$ENSO_{STAR}$	$VECM$	$STVECM$
T	389	389	389	389
m	5	13	132	266
AIC	3.38	3.33	-49.88	-50.39
LLK	-43.80	-35.91	5307.64	5405.34

Table 3: Diagnostic Statistics for the Estimated Models

Figures

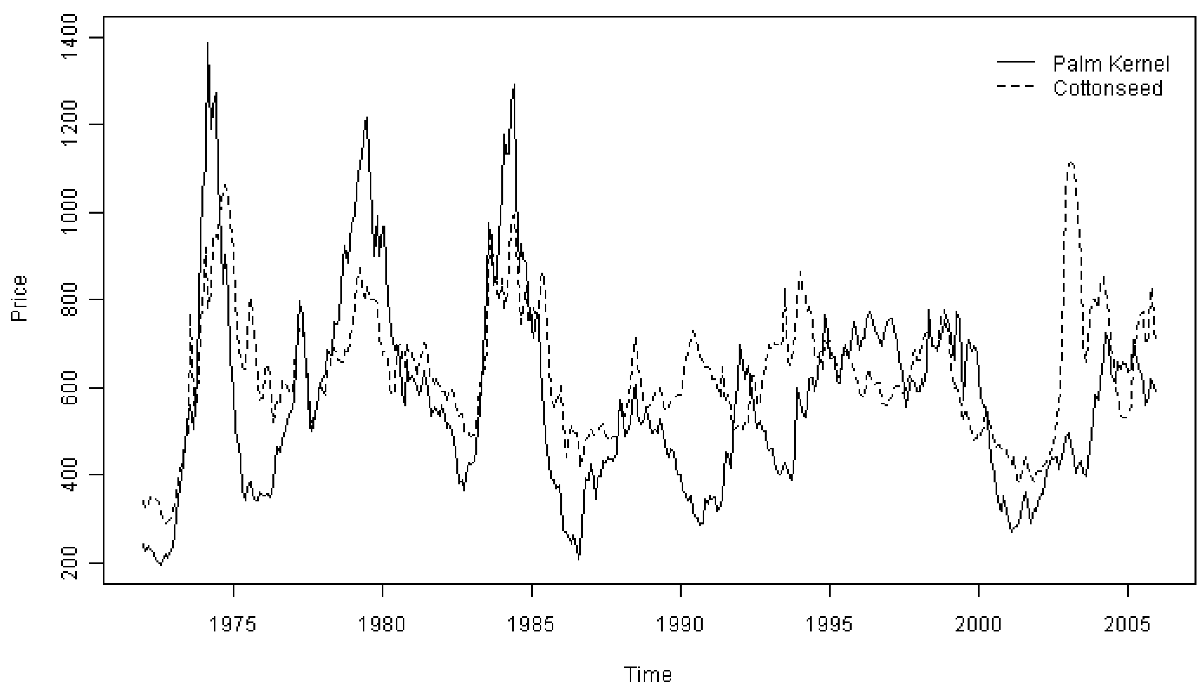

Figure 1: Nominal Palm Kernel and Cottonseed Prices in 1972-2005

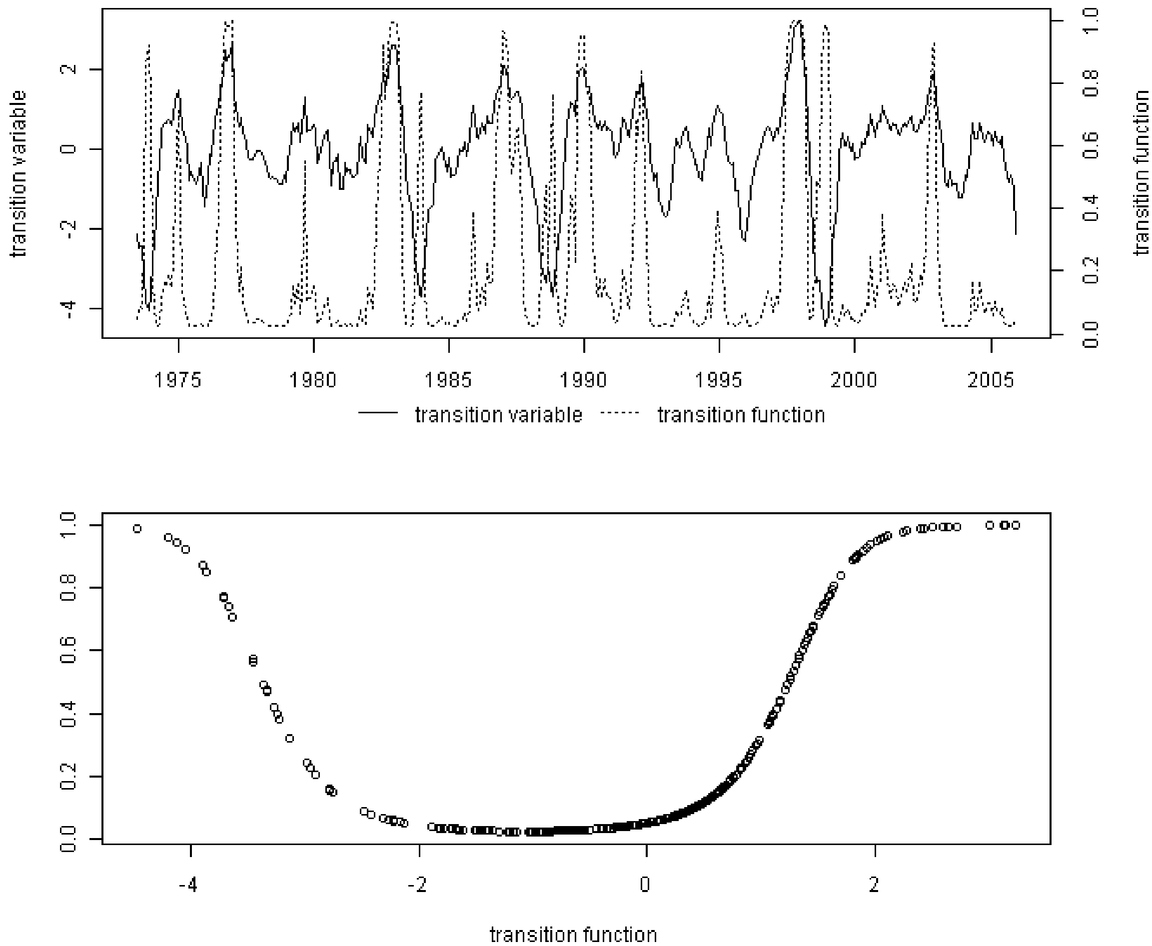

Figure 2: ENSO STAR Model: $\gamma = 1.33$, $c_1 = -3.37^{***}$, and $c_2 = 1.24^{**}$

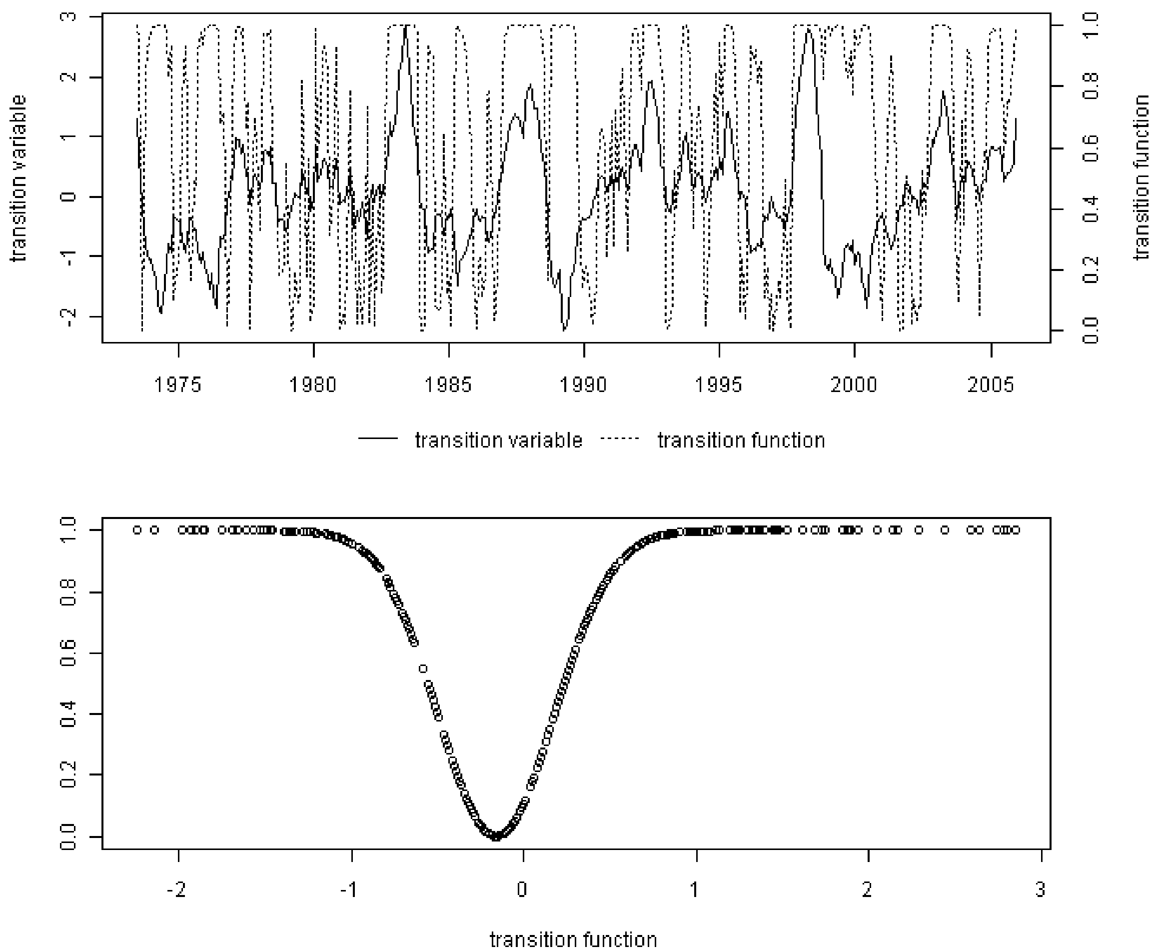

Figure 3: Vegetable Oil Prices STVEC Model: $\gamma = 4.17^{**}$, $c = -0.16$

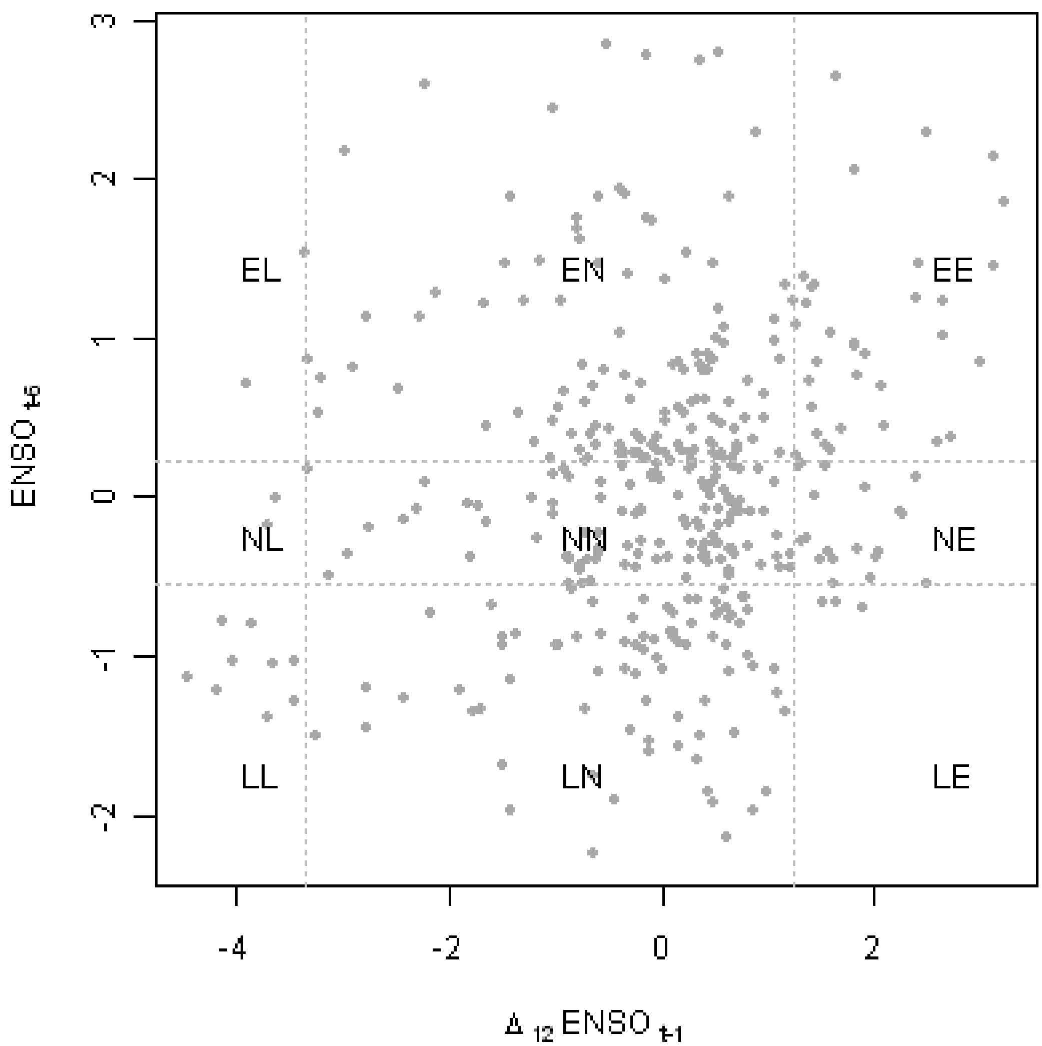

Figure 4: Distribution of the Transition Variables over the Different Regimes

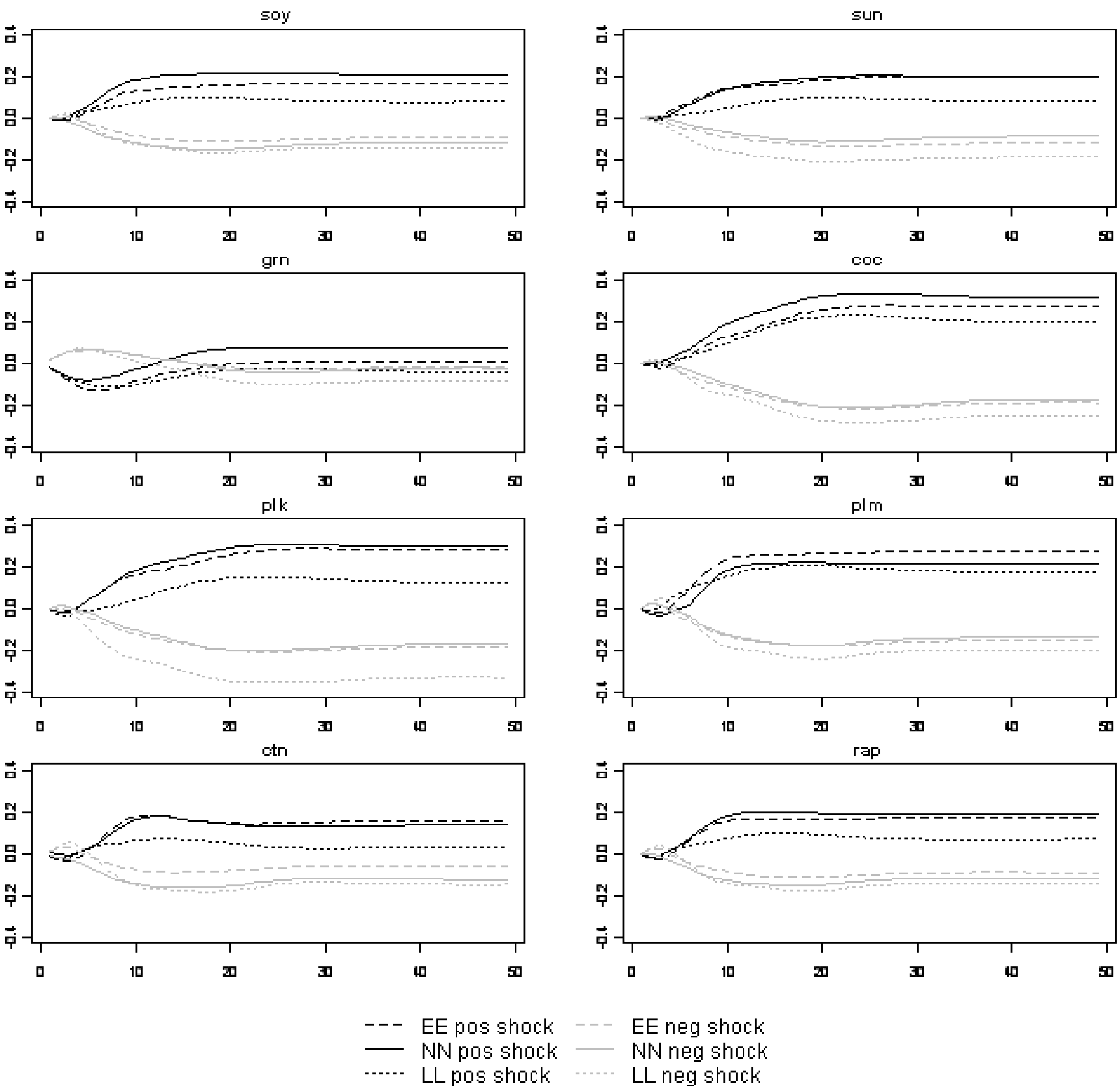

Figure 5: Generalized Impulse-Response Functions